

### LISTING OF THE CLAIMS

Claims 1-30 were originally pending. Please amend the claims 2-6, 8, 10, 12-16, 18, 20, 22-26, 28, and 30. Kindly cancel claims 1, 7, 9, 11, 17, 19, 21, 27, and 29 without prejudice. No claims are added. Accordingly, claims 2-6, 8, 10, 12-16, 18, 20, 22-26, 28, and 30 remain pending.

The following listing of claims replaces all prior versions and listings of claims in the application.

1. (Canceled)

2. (Currently amended) A method as recited in claim 51, wherein the queries comprise a well formed natural language question, a keyword, or a phrase.

3. (Currently amended) A method as recited in claim 51, wherein the query cluster is used to disambiguate a word or phrase in a query of the queries.

4. (Currently amended) A method as recited in claim 51, further comprising determining that the queries are similar based on similar keyword or phrase composition

5. (Currently amended) ~~A method as recited in claim 1, wherein identifying the same document and/or the similar documents further comprises:~~ A method for clustering queries, the method comprising:

identifying a same document and/or a plurality of similar documents selected by a user in response to a plurality of queries, determining the similar

documents being determined by evaluating a set of selected similar documents chosen responsive to queries  $p$  and  $q$  of the queries, wherein documents  $D\_C(.)$  is a subset of a result list  $D(.)$  according to the following:

$$D\_C(p) = \{d_{p1}, d_{p2}, \dots, d_{pi}\} \subseteq D(p)$$

$$D\_C(q) = \{d_{q1}, d_{q2}, \dots, d_{qj}\} \subseteq D(q);$$

wherein similarity based on selection of documents is based on:

If  $D\_C(p) \cap D\_C(q) = \{d_{pq1}, d_{pq2}, \dots, d_{pqk}\} \neq \emptyset$ , then documents

$d_{pq1}, d_{pq2}, \dots, d_{pqk}$  represent a set of common topics of queries  $p$  and  $q$ , and,

whereby the similar documents between queries  $p$  and  $q$  is determined by

$D\_C(p) \cap D\_C(q)$ ; and

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate that the queries are similar independent of whether individual ones of the queries comprise similar composition with respect to other ones of the queries.

6. (Currently amended) A method as recited in claim 54, further comprising constructing a thesaurus comprising a plurality of synsets, wherein each synset comprises one or more query clusters.

7. (Canceled)



10. (Currently amended) ~~A method as recited in claim 9:~~ A method for clustering queries, the method comprising:

identifying a same document and/or a plurality of similar documents selected by a user in response to a plurality of queries, the similar documents being based on a hierarchical positioning between individual ones of a plurality of documents commonly selected across the queries, wherein  $F(d_i, d_j)$  is a lowest common parent node for documents  $d_i$  and  $d_j$ , wherein  $L(x)$  is a level of a node  $x$ , wherein  $L\_Total$  identifies a total number of levels in a hierarchy, and wherein a similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that}$$

$$s(d_i, d_j) = 1; \text{ and } s(d_i, d_j) = 0 \text{ if } F(d_i, d_j) = \text{root; and}$$

~~the method further comprises:~~

incorporating  $s(d_i, d_j)$  into a calculation of query similarity, wherein  $d_i$  ( $1 \leq i \leq m$ ) and  $d_j$  ( $1 \leq j \leq n$ ) be a set of selected documents for queries  $p$  and  $q$  respectively such that:

$$similarity_{hierarchy}(p, q) = \frac{1}{2} \times \left( \frac{\sum_{i=1}^m (\max_{j=1}^n s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^n (\max_{i=1}^m s(d_i, d_j))}{rd(q)} \right);$$

and

responsive to identifying the same document and/or the similar documents, generating a query cluster to indicate that the queries are similar independent of whether individual ones of the queries comprise similar composition with respect to other ones of the queries.

11. (Canceled)

12. (Currently amended) Computer-readable media as recited in claim ~~15~~ 11, wherein the queries comprise a well formed natural language question, a keyword, or a phrase.

13. (Currently amended) Computer-readable media as recited in claim ~~15~~ 11, wherein the query cluster is used to disambiguate a word or phrase in a query of the queries.

14. (Currently amended) Computer-readable media as recited in claim ~~15~~ 11, wherein the computer-executable instructions further comprise instructions for determining that the queries are similar based on similar keyword or phrase composition.

15. (Currently amended) ~~Computer-readable media as recited in claim 11, wherein the instructions for identifying the same document and/or the similar documents further comprise instructions for:~~ Computer-readable media comprising computer-executable instructions for identifying similar queries, the computer-executable instructions comprising instructions for:

identifying a same document and/or a plurality of similar documents selected by a user in response to a plurality of queries, determining the similar documents being determined by evaluating a set of selected similar documents

1 chosen responsive to queries  $p$  and  $q$  of the queries, wherein documents  $D_C(.)$  is  
2 a subset of a result list  $D(.)$  according to the following:

$$3 \quad D_C(p) = \{d_{p1}, d_{p2}, \dots, d_{pi}\} \subseteq D(p)$$

$$4 \quad D_C(q) = \{d_{q1}, d_{q2}, \dots, d_{qj}\} \subseteq D(q);$$

5 wherein similarity based on selection of documents is based on:

6 If  $D_C(p) \cap D_C(q) = \{d_{pq1}, d_{pq2}, \dots, d_{pqk}\} \neq \emptyset$ , then documents  
7  $d_{pq1}, d_{pq2}, \dots, d_{pqk}$  represent a set of common topics of queries  $p$  and  $q$ , and,

8 whereby the similar documents between queries  $p$  and  $q$  is determined by

9  $D_C(p) \cap D_C(q)$ ; and

10 responsive to identifying the same document and/or the similar documents,  
11 generating a query cluster to indicate that the queries are similar independent of  
12 whether individual ones of the queries comprise similar composition with respect  
13 to other ones of the queries.

14  
15 16. (Currently amended) Computer-readable media as recited in claim  
16 15 ~~14~~, wherein the computer-executable instructions further comprise instructions  
17 for constructing a thesaurus comprising a plurality of synsets, wherein each synset  
18 comprises one or more query clusters.

19  
20 17. (Canceled)

1        18.    (Currently amended) ~~Computer-readable media as recited in claim~~  
2    ~~17, wherein the instructions for identifying the same document and/or the similar~~  
3    ~~documents further comprise instructions for~~ Computer-readable media comprising  
4    computer-executable instructions for identifying similar queries, the computer-  
5    executable instructions comprising instructions for:

6        identifying a same document and/or a plurality of similar documents  
7    selected by a user in response to a plurality of queries, the similar documents  
8    being determined ~~determining the similar documents~~ based on a proportionality of  
9    commonly selected individual documents, such that:

$$10 \qquad \text{similarity}_{\text{single\_doc}}(p, q) = \frac{RD(p, q)}{\text{Max}(rd(p), rd(q))},$$

11        wherein  $rd(.)$  is the number of clicked documents for a query of the  
12    queries, and wherein  $RD(p, q)$  is the number of document selections in common;  
13    and

14        responsive to identifying the same document and/or the similar documents,  
15    generating a query cluster to indicate that the queries are similar independent of  
16    whether individual ones of the queries comprise similar composition with respect  
17    to other ones of the queries.

18  
19        19.    (Canceled)

20. ~~(Currently amended) Computer-readable media as recited in claim~~  
19 Computer-readable media comprising computer-executable instructions for  
identifying similar queries, the computer-executable instructions comprising  
instructions for:

identifying a same document and/or a plurality of similar documents  
selected by a user in response to a plurality of queries, the similar documents  
being based on a hierarchical positioning between individual ones of a plurality of  
documents commonly selected across the queries, wherein  $F(d_i, d_j)$  is a lowest  
common parent node for documents  $d_i$  and  $d_j$ , wherein  $L(x)$  is a level of a node  $x$ ,  
wherein  $L\_Total$  identifies a total number of levels in a hierarchy, and wherein a  
similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that}$$

$$s(d_i, d_i) = 1; \text{ and } s(d_i, d_j) = 0 \text{ if } F(d_i, d_j) = \text{root}; \text{ and}$$

~~wherein the computer-executable instructions further comprise~~  
instructions for:

incorporating  $s(d_i, d_j)$  into a calculation of query similarity, wherein  $d_i$  ( $1 \leq i \leq m$ ) and  $d_j$  ( $1 \leq j \leq n$ ) be a set of selected documents for queries  $p$  and  $q$  respectively such that:

$$\text{similarity}_{\text{hierarchy}}(p, q) = \frac{1}{2} \times \left( \frac{\sum_{i=1}^m (\max_{j=1}^n s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^n (\max_{i=1}^m s(d_i, d_j))}{rd(q)} \right); \text{ and}$$

responsive to identifying the same document and/or the similar documents,  
generating a query cluster to indicate that the queries are similar independent of



1 whether individual ones of the queries comprise similar composition with respect  
2 to other ones of the queries.

3  
4 21. (Canceled)

5  
6 22. (Currently amended) A computing device as recited in claim 25 21,  
7 wherein the queries comprise a well formed natural language question, a keyword,  
8 or a phrase.

9  
10 23. (Currently amended) A computing device as recited in claim 25 21,  
11 wherein the query cluster is used to disambiguate a word or phrase in a query of  
12 the queries.

13  
14 24. (Currently amended) A computing device as recited in claim 25 21,  
15 wherein the computer-executable instructions further comprise instructions for  
16 determining that the queries are similar based on similar keyword or phrase  
17 composition.

18  
19 25. (Original) ~~A computing device as recited in claim 21, wherein the~~  
20 ~~instructions for identifying the same document and/or the similar documents~~  
21 ~~further comprise instructions for~~ A computing device comprising:

22 a processor coupled to a memory, the memory comprising computer  
23 executable instructions, the processor being configured to fetch and execute the  
24 computer-executable instructions for:

1 identifying a same document and/or a plurality of similar documents  
2 selected by a user in response to a plurality of queries, determining the similar  
3 documents being determined by evaluating a set of selected similar documents  
4 chosen responsive to queries  $p$  and  $q$  of the queries, wherein documents  $D_C(.)$  is  
5 a subset of a result list  $D(.)$  according to the following:

$$D_C(p) = \{d_{p1}, d_{p2}, \dots, d_{pi}\} \subseteq D(p)$$

$$D_C(q) = \{d_{q1}, d_{q2}, \dots, d_{qj}\} \subseteq D(q);$$

8 wherein similarity based on selection of documents is based on:

9 If  $D_C(p) \cap D_C(q) = \{d_{pq1}, d_{pq2}, \dots, d_{pqk}\} \neq \emptyset$ , then documents  
10  $d_{pq1}, d_{pq2}, \dots, d_{pqk}$  represent a set of common topics of queries  $p$  and  $q$ , and,

11 whereby the similar documents between queries  $p$  and  $q$  is determined by  
12  $D_C(p) \cap D_C(q)$ ; and

13 responsive to identifying the same document and/or the similar documents,  
14 generating a query cluster to indicate that the queries are similar independent of  
15 whether individual ones of the queries comprise similar composition with respect  
16 to other ones of the queries.

17  
18 26. (Currently amended) A computing device as recited in claim 25 ~~24~~,  
19 wherein the computer-executable instructions further comprise instructions for  
20 constructing a thesaurus comprising a plurality of synsets, wherein each synset  
21 comprises one or more query clusters.

22  
23 27. (Canceled)  
24  
25

1       28. (Currently amended) A computing device comprising as recited in  
2 claim 27, wherein the instructions for identifying the same document and/or the  
3 similar documents further comprise instructions for:

4       a processor coupled to a memory, the memory comprising computer  
5 executable instructions, the processor being configured to fetch and execute the  
6 computer-executable instructions for:

7       identifying a same document and/or a plurality of similar documents  
8 selected by a user in response to a plurality of queries, the similar documents  
9 being determined determining the similar documents based on a proportionality of  
10 commonly selected individual documents, such that:

$$11 \quad \text{similarity}_{\text{single\_doc}}(p, q) = \frac{RD(p, q)}{\text{Max}(rd(p), rd(q))},$$

12       wherein  $rd(.)$  is the number of clicked documents for a query of the  
13 queries, and wherein  $RD(p, q)$  is the number of document selections in common;  
14 and

15       responsive to identifying the same document and/or the similar  
16 documents, generating a query cluster to indicate that the queries are similar  
17 independent of whether individual ones of the queries comprise similar  
18 composition with respect to other ones of the queries.

19  
20       29. (Canceled)  
21  
22  
23  
24  
25

30. (Currently amended) A computing device comprising as recited in claim 29:

a processor coupled to a memory, the memory comprising computer executable instructions, the processor being configured to fetch and execute the computer-executable instructions for:

identifying a same document and/or a plurality of similar documents selected by a user in response to a plurality of queries, the similar documents being based on a hierarchical positioning between individual ones of a plurality of documents commonly selected across the queries, wherein  $F(d_i, d_j)$  is a lowest common parent node for documents  $d_i$  and  $d_j$ , wherein  $L(x)$  is a level of a node  $x$ , wherein  $L\_Total$  identifies a total number of levels in a hierarchy, and wherein a similarity between two documents is defined as follows:

$$s(d_i, d_j) = \frac{L(F(d_i, d_j)) - 1}{L\_Total - 1}, \text{ such that}$$

$$s(d_i, d_j) = 1; \text{ and } s(d_i, d_j) = 0 \text{ if } F(d_i, d_j) = \text{root; and}$$

~~wherein the computer executable instructions further comprise instructions for:~~

incorporating  $s(d_i, d_j)$  into a calculation of query similarity, wherein  $d_i$  ( $1 \leq i \leq m$ ) and  $d_j$  ( $1 \leq j \leq n$ ) be a set of selected documents for queries  $p$  and  $q$  respectively such that:

$$similarity_{hierarchy}(p, q) = \frac{1}{2} \times \left( \frac{\sum_{i=1}^m (\max_{j=1}^n s(d_i, d_j))}{rd(p)} + \frac{\sum_{j=1}^n (\max_{i=1}^m s(d_i, d_j))}{rd(q)} \right)$$

and

1           responsive to identifying the same document and/or the similar  
2 documents, generating a query cluster to indicate that the queries are similar  
3 independent of whether individual ones of the queries comprise similar  
4 composition with respect to other ones of the queries.  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25